



A DOCPHOENIX

# Raw Sequence Listing Error Summary

## ERROR DETECTED SUGGESTED CORRECTION

SERIAL NUMBER:

09/701,626

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

1      Wrapped Nucleics

The number/text at the end of each line "wrapped" down to the next line.  
This may occur if your file was retrieved in a word processor after creating it.  
Please adjust your right margin to .3, as this will prevent "wrapping".

2      Wrapped Aminos

The amino acid number/text at the end of each line "wrapped" down to the next line.  
This may occur if your file was retrieved in a word processor after creating it.  
Please adjust your right margin to .3, as this will prevent "wrapping".

3      Incorrect Line Length

The rules require that a line not exceed 72 characters in length. This includes spaces.

4      Misaligned Amino Acid Numbering

The numbering under each 5th amino acid is misaligned. This may be caused by the use of tabs between the numbering. It is recommended to delete any tabs and use spacing between the numbers.

5      Non-ASCII

This file was not saved in ASCII (DOS) text, as required by the Sequence Rules.  
Please ensure your subsequent submission is saved in ASCII text so that it can be processed.

6      Variable Length

Sequence(s)      contain n's or Xaa's which represented more than one residue.  
As per the rules, each n or Xaa can only represent a single residue.  
Please present the maximum number of each residue having variable length and indicate in the (ix) feature section that some may be missing.

7      PatentIn ver. 2.0 "bug"

A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequence(s)     . Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies primarily to the mandatory <220>-<223> sections for Artificial or Unknown sequences.

8      Skipped Sequences (OLD RULES)

Sequence(s)      missing. If intentional, please use the following format for each skipped sequence:  
(2) INFORMATION FOR SEQ ID NO:X:  
(i) SEQUENCE CHARACTERISTICS:(Do not insert any headings under "SEQUENCE CHARACTERISTICS")  
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X:  
This sequence is intentionally skipped

Please also adjust the "(iii) NUMBER OF SEQUENCES:" response to include the skipped sequence(s).

9      Skipped Sequences (NEW RULES)

Sequence(s)      missing. If intentional, please use the following format for each skipped sequence.  
<210> sequence id number  
<400> sequence id number  
000

10      Use of n's or Xaa's (NEW RULES)

Use of n's and/or Xaa's have been detected in the Sequence Listing.  
Use of <220> to <223> is MANDATORY if n's or Xaa's are present.  
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.

11      Use of <213>Organism (NEW RULES)

Sequence(s)      are missing this mandatory field or its response.

12      Use of <220>Feature (NEW RULES)

Sequence(s)      are missing the <220>Feature and associated headings.  
Use of <220> to <223> is MANDATORY if <213>ORGANISM is "Artificial" or "Unknown"  
Please explain source of genetic material in <220> to <223> section.  
(See "Federal Register," 6/01/98, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of new Rules)

13      PatentIn ver. 2.0 "bug"

Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing).  
Instead, please use "File Manager" or any other means to copy file to floppy disk.

PCT

## RAW SEQUENCE LISTING

DATE: 01/30/2001

PATENT APPLICATION: US/09/701,626

TIME: 11:15:33

Input Set : A:\Neb-165a.app

Output Set: N:\CRF3\01302001\I701626.raw

Does Not Comply  
Corrected Diskette Needed

3 <110> APPLICANT: New England Biolabs, Inc.  
 4 Vaisvila, Romualdas  
 5 Morgan, Richard D.  
 6 Raleigh, Elisabeth  
 8 <120> TITLE OF INVENTION: Restriction Enzyme Gene Discovery Method  
 10 <130> FILE REFERENCE: Gene Discovery Method  
 12 <140> CURRENT APPLICATION NUMBER: US/09/701,626  
 13 <141> CURRENT FILING DATE: 2000-01-12  
 15 <150> PRIOR APPLICATION NUMBER: 60/089,086  
 16 <151> PRIOR FILING DATE: 1998-06-12  
 18 <150> PRIOR APPLICATION NUMBER: 60/089,101  
 19 <151> PRIOR FILING DATE: 1998-06-12  
 21 <160> NUMBER OF SEQ ID NOS: 130  
 23 <170> SOFTWARE: PatentIn Ver. 2.0  
 25 <210> SEQ ID NO: 1  
 26 <211> LENGTH: 14143  
 27 <212> TYPE: DNA  
 28 <213> ORGANISM: Unknown  
 30 <220> FEATURE:  
 31 <223> OTHER INFORMATION: Genomic DNA of Pseudomonas Alcaligenes NEB#585  
 32 (ATCC 55044)  
 34 <400> SEQUENCE: 1  
 35 atcgatcagc cagacttttc gcacacgggc ggaccttggg cgagtcagcg ctatgggttg 60  
 36 ccgctgtggg ttgtcagtcg ccgtacgcgc aatctgtttc ttctgcaggg catgtccggc 120  
 37 tgggcgttcc ggcccgttct ggtcaccgac tcggtctctc atgagcgcta tctcgctcta 180  
 38 agtcaggaac ttgcgcaact gcttcgtgat gcaccgcaga gcaagctcga agaccgtgat 240  
 39 tggtaagcgg gggctattcg atcagtcctg gaggcgacaa actccagaaa cgacaaggcc 300  
 40 ctgaaaaaaa agcagggtt cgtcttttgc ggcgaaatga atcggaacctc ttcccgctc 360  
 41 tgcattgtaac tggcttttgt ttgcccatac tgcctatctc atgccggcca tgttggccag 420  
 42 tgcctgcata atttggcctt tgggttgcac actttttcga cagccctgct agacatccct 480  
 43 cctctgccc tcgttaacttc tgttccgatg gtgtcgcttg gcactatggt cttgtcgagt 540  
 44 gtccgttttc atccagccta atgcccgcat tgcctcgctg agctgtagct gaatacagga 600  
 45 cttagcggac gacaagggaat gttatgcgaa acatgtlggc gaataaatla cggcgcatgt 660  
 46 ttcgtctact tatagttagg ctacatatga gaatcagcgc agaccagctt gctcaagaa 720  
 47 cactgactga gttcggcgctg ctggcggtta agcttctggc aacycgagag cttagccagt 780  
 48 tgtccgaqaa gtttgggtat gcactggcct tcggaaggga accggcggtt gccatagctg 840  
 49 aggaccttgc taggtgcttg tgcggacaaa atgcttcgcc ggcatctgaa taccocaaaa 900  
 50 tcaccgttaa gtatttcaag gaaaaagaaa gtagtctggt ggcactcgta gagtgttatg 960  
 51 tacaaatgac cgcaagcgca aacattcttt tagagctggt tgcgcacga aatggagagg 1020  
 52 caataaatct gtatctagaa ggccttgagtg ttgtlagicc acaatgcgct caaagcgctc 1080  
 53 acttcgttgc ctgggaccgg cgaagccggc ccttagctt aatcgttaga aaccatcatg 1140  
 54 gataactggt acaacacat cgaataccac acccatgtag ccgaaaaact agaggcaatl 1200  
 55 ggagaaacaa agtacgaccg cgaggcttat gaattcgcgc tagaggcata ccagtatgcy 1260  
 56 cctgaatata atgaaaatat tcccacgcgc cctctcaatc ttgggctcgc gtaccatgta 1320  
 57 agcgcttca actttgcaca ctgctatgta cttcaccgta aagaagtgtt tgaagctcca 1380  
 58 aaagacacac tgagctcctg gggcgatttt tccccaacgg acattggtga aattgtttat 1440  
 59 ggttttagtcc gtattggctt gctggaccaa ggcgccgaag acaaaaaaga gcagtttgaa 1500

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/701,626

DATE: 01/30/2001

TIME: 11:15:33

Input Set : A:\Neb-165a.app

Output Set: N:\CRF3\01302001\I701626.raw

```

60 ggggttgtttt taatcaccca cgtgctgtga tgtcttctaa ctactgggttc aagtcgttcg 1560
61 cttcgctcac tcgggaccgg ctaaagccgg ccccttaacc aaacgttagc caccctcacga 1620
62 aqattttggag cccgcgtgaa caaagctgat acaaaacaaa ttaaaacgga tttttcgga 1680
63 cgaattgatg aaaaaagagc gtggtttgat cgtatggcta cgtttataag cgggacaaac 1740
64 accgagttaa ccgaccttaa tttcttttgc gagaactata taacatcaat atacgtagag 1800
65 ctogaatgct taataicaga tttatttcat ggctacataa ataacaacaa caagacctac 1860
66 atggcgacaca ttcaatcaaa aatcaagaac tccataactg acaagtactc tgcattggcac 1920
67 gccacccata caacattgc aggtccagay catattaatt cagcacagct cagcacgctc 1980
68 cttgatccaa caagctggaa catcacattt aaagacgttt ccgcaatgaa agtacgagca 2040
69 aaggaaatacc tttctctagt acacgaaaaa agattttcag gtatatctgc atccgatgga 2100
70 gctcttattg atgcgcgaca tgcattcaga aattgcattg cacacaacag cgaagctcc 2160
71 aaaaaggtta tgaacaccaa aattaaaaag ttaattacag gccagcttg ctaaatgtc 2220
72 ggccttgaac tcaccacaaa tagtgtgacc aaaataggaa agtatctccg tgcattgct 2280
73 cagcaaaaga tgcgagtgct gatttactca gatcgataa aatctatcgg ctaagctta 2340
74 taagtgtggg ctaacaatgc gctcaactgt cgtcacttcc gttcgctgga cagccaaaag 2400
75 ctacgctttt gtcgcccgtg tagcttaatc gttaggaagg tctgcatgac tctgcaaca 2460
76 gacaggttcc aagagcttct gcaatcacat gagttctcag ggcattattat tctgtgggtt 2520
77 gcgatattcg aaggccgtct tgacgggtgtg ttatcagttc attttcttgg acttgaaagc 2580
78 acctatgaat tctacgaact catactttcc aggttgtctt tctacgaaaa aattgaaatc 2640
79 ctgagaaaaa ttgattttgg taacagtctc aaatcccaag aaaatacagc gctgcacctc 2700
80 gcaaaactga ggcgatttgc taacgcattg gcgcattgag cacacatgcc acctgatgaa 2760
81 atcatgaaat tgtgctctga taagtggata gagtctcttg tgcctggata tccaaagtcc 2820
82 attggcaaaag agaaaaatgc acttgaaaat cggctatcac ttctgtggaa ttactgccac 2880
83 aggaagcatg tagcaaaaat taagcagctt gcacacgaac tcaaaaatac agagcaagcc 2940
84 aactaataga gtccagttat acaggtccgt aaatgagccg cctaacaact ggttcaagcc 3000
85 actcaacttcg ctgcgtcggg accgcttcc gcggccctt aaccaaactg tgggcaccca 3060
86 tagaaaaatc ctaatgagaa aactattcat arcactaatt ttgcctctgc tctcgagag 3120
87 cttgatggca tctgaagcgt ataaggacct tgaaacacaa gtaactgaaa aagccagcct 3180
88 agcagttgcc caaatgaatg acagagcaac tggaaaagctc gactactcgg aagaaagtct 3240
89 ctatgcagta gaagaaatgg cagcggaaag agctcaatc aaagatcaat tagatccagc 3300
90 cactgtagac tgccttactc aagttcttgg aagctatatt cttgaggttg cactagaaaa 3360
91 gcatggcgcg tcttaagttt gqcttgaatc tgaaaaactc cctgccttgg taqttggtg 3420
92 accagagtac aggttagcac tctcaacctt cgcgaaggta catggccgac tttctggcga 3480
93 cgaagcagat aatcttattt tcttctatca aggttttctt gaaaggctta aatcaccatc 3540
94 tcccgccatg agcgcactct acaaatgaaa cccgagtttg ggggcccaac aatgcgctca 3600
95 actgcgcctc acttgccttg ctggaagctc aaaaagctac cttttggccg cccgttagct 3660
96 taatcgttat gcacaataaa acatgaagac agcactcata tttgtagctc taatctttct 3720
97 ctctggtagt gacaactatc agtcatgccc tataactgga aaatggaaat ccaacgaaaa 3780
98 gctaaactta gaaagcatga atgaaaacgg caggataaac gcaaaagcaa gagagatttt 3840
99 tgagaacggc ttcttttgaa aactagaatt agacataaat tgcagtagct tcacaacaat 3900
100 acttgacggc gttaccgaaa cctttaattc cgagatagtt cgcacaaaca aagattccgt 3960
101 caccgttagc tattacagca aagcgtgca aaaacaagtt gaggtcacat ctattatcaa 4020
102 cggaaaattgt tactcgacac ctatagagca gttaaatttc aatqagtatt tctgcagaat 4080
103 cgagtagcgc ataacaattg attcaagtgc ttgccttgcg tcaactgggg accggctaaa 4140
104 gccggccctc taaccaaagc ttaggcaaa gctcaatgga tcccatattc cataacatcc 4200
105 atagaaacga caaagagatt gagggcgctc atcaacaatg ctcgagcaca atcaatcact 4260
106 tcattgagat ggtcaaaaaa gggggcgagc ccacctatat ggcaaaagct cgttttcttg 4320
107 accctgacaa gtctgaaaaa gaaggttaaga atcatatttt ttatttggg ttatctgaag 4380
108 tctgttacca cctgcaaca aatttacttt ctggggtatt ttttgaatc cctgaaggt 4440

```

## RAW SEQUENCE LISTING

DATE: 01/30/2001

PATENT APPLICATION: US/09/701,626

TIME: 11:15:33

Input Set : A:\Neb-165a.app

Output Set: N:\CRF3\01302001\I701626.raw

```

109 ttgaaaagtg gcaccaaata ggccagcgcc taggctttga tccagaagat gtctttgatt 4500
110 ggatggtaat cgacaaaagt catgctaagg gtgcatacac actaaaggta tcgcgagagc 4560
111 gcttaaccac cgagcaagaa agaaaagatt ttgaccgcta tattgggtgt gcgtcatatg 4620
112 agtagcctaa aattaagcgc tcacgcctca qcctaactac tggttcaagt cactcgcttc 4680
113 gctcgttcgg gacgcggttc cgcggccct taaccaaacg ttaggcgcaa gggcaatatt 4740
114 ggtcttcagc accgagtcag gaaacacaat caccgaatca gcgcggtgtt cctgaatcga 4800
115 atggtcgtcg acagttagg cgtttatttg tggccagcaa aggagttgct ttcagagaat 4860
116 gtgcacgtca caaataactt ccggggccaa aaccgaaacg ccgtgcgctc cgcgggttaa 4920
117 gctcggcgct ggcatcattt tcggcgctcg gctgggcaat ctaacaattg gctcaagtcg 4980
118 ttgcgttcgc tcactcggga ccggcggaag ccggccctta gccaaacgtt atgcgagcca 5040
119 ccatgaatag cgaagaatta tacaaaaagg ctatggaggt agagtccaaa tgcgagcata 5100
120 aagcgccaat ttcaacttac aagaatttg ttaagaaatc taacgatcct cgacacttca 5160
121 tcgcattcgg agtttgcttc caaaaatgtg gtcactggaa gcaatccatc qaggtattag 5220
122 aatcaggaat tgcactgaag cctcactatt gcgaggggtg tgcctgtcta tttttagcaa 5280
123 aagcactttt taatcaggc aaaaaaggcc tlgcgataaa gcaatggcaa catglatcaa 5340
124 aaatgcaacc tgagtaccca agttatgagt ctgtgcacaa tgaagccaaq aaantgcttq 5400
125 cacaaaacgc ataaccaatt gctcaagccg ctgcctccgc tcaactcgga qtcctgaagc 5460
126 tacgcttcgg gcgcgccctt agccaaacgt taggcaccac atgcctccca tcaagtcagc 5520
127 aagcraatac cagcgcgcaa catcgctcat cttcttagtg tcaggcgctg cttgctattt 5580
128 catcgtgcag tcttcgttgc tgccttgac ggtatgtgcc cgcacaaqaa tggtttgctt 5640
129 taatategtt cttgggtatt cctgttttgt tataggtagt gcgcgcaaag gtcagcgaga 5700
130 atttcgctgc cctgactgtg ggaacgaagt agatcagagc ttacctacag agggtgatgg 5760
131 cgcceccact ctaaggctgt gcaagcactg cgatattcta tggaaatgtt gcaagacccc 5820
132 agacagltta agttaccgcc taacaaatgg ttcaagtctg tcgcttcgct cactcgggac 5880
133 cggctaaaag cgcgccctta accaaacgtt aggcaacagg gggtagatg acgcaatgtc 5940
134 caaggtgcca gcgcaatctc gcagctgacg agttctatgc tggctctagc aasatgtgca 6000
135 aggggtgcat gacttgccaa aacctaaagt acaacgcgaa taagggaagt catgccaaca 6060
136 ccttcaccaa agcagacatt ttggcgttgt acggttato aqcacagcgg cattgtgggt 6120
137 attgcgggat atcggaggca ggttttacat ccttgcacag gactaatcca cgcggctacc 6180
138 acatacagtg tttgggtggt gatcgctcag attcgttcga aggtatttca cctcaaaacg 6240
139 ctgcgctcgc ctgttttata tgcaacagga taaaaatcaa catcttcagc gccagtgaqa 6300
140 tggagcttct aggtgagggc atttcaaaaq cgtggcatgg tcgaggaatt qcctaactat 6360
141 tcagtcaagc ggaagcaaac cccgctgcgc ggtctttgcg ccgcttatct caagcgttag 6420
142 atgaataaaa gctccacac atagccagct ttacgggaac gaagttgatg cgaagcctct 6480
143 ttactgcctt agtattaaat ttgatgattt ctgggtgtct ctcacatca aaaactgaaa 6540
144 gccataaaca gcgcgcgaac aattcaagcg acacgaccgc catactgaaa tatatattta 6600
145 ctgttaccca tggcatagaa gagggtttg ocaagtcact tgagcaagga agttacacac 6660
146 cagatgaata tatgctcatg cagcaggctt tttagcaatc tgatctaaac agattaaaca 6720
147 ctctttttat acccacttta gataaaaagc tgaacaccgc agacgtaaaa catttcatga 6780
148 tttttataaa atctactgca ggtaaqaact tgctaacagc aggagagtct agcaactctt 6840
149 tttccggcac aatggataga gttcgatcac taccctccga acagcagtcg aaaataaacg 6900
150 agttcttcca tgcagcttac acaaaaaaca ctttaacagc catgggggct ccagagggcg 6960
151 taagcattgt ttatgcattt gtagtggaat ccatgtgcaa ttacgctatg cgcataaatt 7020
152 tcgaactgta tttttctatt atagaaaaag gcaaatgcca ataaaccaca aaaaacaaag 7080
153 gccgagtcga tccatgtgaa catcagcgtt acatctaaca tgtgtttcaa gccgctcgct 7140
154 tcqctcactc ggaacgggct aaagccggcc ccttaaccaa acgttagagg attacatgcc 7200
155 atcactccaa gaactccaat cgcgatttga ctacgaatc gtgaactcca tgattgagag 7260
156 cactcccgaa acatgaagcc agattattct aacgttggtg cgcgaatcta attcctcttg 7320
157 tgtaggtaac tttacacatg agttatccag tccgtagggg catgcaccag ttggtccgpc 7380

```

## RAW SEQUENCE LISTING

DATE: 01/30/2001

PATENT APPLICATION: US/09/701,626

TIME: 11:15:33

Input Set : A:\Neb-165a.app

Output Set: N:\CRF3\01302001\I701626.raw

```

158 agagagctta tttgaagaca cttaccaact cgtagagtta ttctacagcc atgggtgagcg 7440
159 cttattcacc aaggcaattt atcgggctaa tgcgggtggg gacgggttgg catabcagc 7500
160 tgagtttgaa tatgcgtagc atccctctaa caattggttc aagcgcctcg cttegcctgc 7560
161 tcgggatcgq cgaagccggc accttaacca aacgttagag atggtcatga ataaacgcgc 7620
162 actaaccttc gggctactca tagcaattct agctagtatc atttcacaag cgtgcttca 7680
163 tggccaaaag gtaategett ccgacgtggc ttatatact ccatattlcc tatctccaat 7740
164 atttaqcatg ctgatacaag ccacatcgat gctgqcatgg qccataacct gactltatgt 7800
165 aggtacttta tgcaagaaca agccagcaca acatggagca aaattggggg cagcatatgy 7860
166 aatacttctt ggatttaatt tattegcaat gcgagcttcg acccaattaa cgtagattct 7920
167 aagttaataa tcgcaacatc tgccttaaca caaaaagcaa aataltcagt gcactttgcg 7980
168 ctagtgtctc ctgcgggcta tcttcttgca aagcatcgtg caaatctcta acaattggtt 8040
169 cagatcggtc gcttcgctca ctgcgggacc ggcggaagcc ggcctcttaa ccaaacgtta 8100
170 ggcaactgaa tgatcacctg cattccggca cgtgaattcc tgcgtaaagt atgcggcctg 8160
171 tacgaagcct caactaatgt agttaagttg cgtgtatggg ctgtgtgata tggcatcgca 8220
172 atggatctaa ctgtcaaaqy taaatctgtc ctttgtgcgg ttgcgggagt actccgcca 8280
173 aggtcgaat getttgctca aatttgctt cccaacgtaa ttcaattagt aggcgacaag 8340
174 cgttcagaga atcaactaaa gctcataggg atggaaccac caatcgaaat tcatactct 8400
175 cgcgaacaga gcaggctcca agttgtaatc ttgtacaggg gtcaggtaaa ggtcacatat 8460
176 gtgctgtcag ccgcttaact actgggtcaa gtcgttcgct tcgctcactc gggaacggct 8520
177 aaagccggcc ccttaacca acgttaggct ttcaatgaaa acagttccag tgaataatc 8580
178 agaagtcgaa ctaatagaga gttttgggaa attcctgac aatcaagact taatcgacta 8640
179 tgaaaaatcc cacttcagt gcgacgacaa ccaaatgca gatgtagct tatctttaa 8700
180 gccagggaaa tggccaggca ttcaagtcca taaactacac atagaagtaa agtcacacca 8760
181 ctcaagaagc tctcaaaaca ccatcaacaa aatattcgcc caatlactaa aagaacccg 8820
182 aaagcgaagc ctogataaag agaaagagtg cttagctata ttgttccctt acgagcgcgg 8880
183 cgcattggca ggtcgaaaca acaaaacagt aacaagaatt gaaggtgaag cttattaccg 8940
184 gaggggcttt tcgagaatcg acaaacagac gtttgtttaa ttgtgtgact tggctcggtc 9000
185 caaatcacat ctttctcttt ctacagcatc aaacacattg aacgtatttt aatggaaaaa 9060
186 tttcttagat gaggaaatca gcccgatgat cagcctaaca aatgggtcaa gccgttcgct 9120
187 tcgctcactc gggacccggc aaagccggcc ccttaacca acgttagagc caccggaaat 9180
188 ttgtcatggg aaaccagaaa atggatttgc agataaacga tacaaggtt vagtgggtt 9240
189 ctccaatact gaagcaatgg atcagcatca acaagaata cgtcaagcaa tatgatttca 9300
190 aagactgcct gcactggtat aacgaaagg gcaatataag tgcctttgct ggtgcggtt 9360
191 ggaagtctgg aggttttgcg ctggaagaat attcaactaa aaaagggacc gaagaaaaca 9420
192 gagccaatgg tcgtgtcgac ctatatctct ccaatgacaa cgaagcaagc attgttgaag 9480
193 caaaaatgga atggctctac ttcgaaagc gcacaagact agatttcaa gaaaaaatag 9540
194 atcgtgtagt tgaaaaagca aagaatgaca taattaacag cctgcatgcc aaccctacg 9600
195 atctagggct tgggctttcc ttattttgca catactggaa aaagggttat gacgcatecg 9660
196 ccgacatgca agcccttaga gcgcttatgc aaaattataa ctgcgcattt tatgcaattt 9720
197 ttgaaaaacg ccccgacaac gaaattgtta gctcaaaagg caatalctgc aacgctgtga 9780
198 ttttagttgg gacggcgcac agctgaatcg tgtgtgtgog totaacaatg cgtcacaagc 9840
199 gctcacttgc ttgcgtggga tcggctaaag ccggccctt agcttaatcg ttgacactag 9900
200 gacttccgac catcatgagt gatagagacg aattttctgc cccaacaaaa agaagcctag 9960
201 ccgaaaaggag tggcttttag tgttcttata ttggttgcct taatgcaacc atagggccta 10020
202 gtgaagaatc agaaacagcc gtacgaagaa cgggggtggc gtgtcatata actgcccgag 10080
203 cgcccgccgg aaaaaggatg gacccaacat taagccctac ggaacgaagc tcaatctcga 10140
204 atggtatatg gatgtgcca acgcattcag ttgaaataga tagagatgag gcccgataca 10200
205 categaacct attaaatcac tggaaaaata talccgagag ccgagcagat tatgcaaaaa 10260
206 atcatggctg gytatatttt gacaaatacc ccttccctca tattgactcg ctagccaaca 10320

```

RAW SEQUENCE LISTING  
 PATENT APPLICATION: US/09/701,626

DATE: 01/30/2001  
 TIME: 11:15:33

Input Set : A:\Neb-165a.app  
 Output Set: N:\CRF3\01302001\I701626.raw

```

207 tagacctggc tcttaccaaa agcccttccct caaatagcct tatcggggaat gccattacag 10380
208 acagctgccc cccccaacta tggggtaaaag agcaatctgt aatcaccaga gacctataaa 10440
209 tagaacttta tcgaaatgcc ttcgatcacg gcgaggtctg ctcattcgaa atatccatat 10500
210 cggagcaaaa actagaaata gtttacgatg gcaaaaaatt tgacatcttc caactctctt 10560
211 accaccagaa tgcaaacggg ggcgcgcgata ccttgcaaga aattgtagaa aaatatggca 10620
212 gtaacttttg agtcaactat agccacgaag gcaacaataa aataataatt cacaggctct 10680
213 ctgactttta cgcgcttgca ccateccctc cgtgcgtaat atcactgaat gaatacagtg 10740
214 acaagqccc ctaggttagac ctggctatct atgagcgctg cggcgactg tacataattc 10800
215 taccggttga tttttgtaga tcagatgtca gggggctaga gtcgcagcta gccgcctttg 10860
216 aacctaatgg aaagccagtt tacattgtag gctcagatgt ggcagagcct acaagaaaag 10920
217 caattataga caggcttccc aacttcacgl tctccaaaaa gcaatgctaa caatgcgctt 10980
218 aactgtcgct cacttccttc gctgagatgt caaaagctgc gcttttgtct gtccgttagc 11040
219 ttaatcgtta ggcgcaagga ctgaaactga gaaatgggtg ggttaagttcg 11100
220 tcagcgggtt tggcgggcag agataccgag aaatttttga agtccctgaa tccagtaacc 11160
221 ttcgcccact gggcaagtc aatactgaaa cattgctatt tcagcttcga ggggctgata 11220
222 qtqaaatgct agatattttt gcccttgcct tggggcgcgc gccagtaatt tcgtttccca 11280
223 aatcatattg gctaggtcgc cccagtgaaat taagcgctca tctatccaat tttcattctt 11340
224 cggaaaagcc agccataaca ggcgcggttt ctgactcaca gtattcgcca ggcaggtg 11400
225 aatcaccgc cctactcat gagaggatta ttgaggtttg caaccgtgct tgtgcttccc 11460
226 tgcaataagc gccataacaac tggttcaaat cgtcgcctcc gctcgcctgg accgcctatg 11520
227 cgggcctctt aaccaagcgt tagatgcaaa taacttgagg ggcacatgca agactttggg 11580
228 tccagacgaa atgcactcat agaggacagg gctgcggctg agtctgttat tgaacgtgtt 11640
229 tatcttgcca tacagcagct ttgcacagag actggtgacg taagaaatcg gcttcaataa 11700
230 gccgttatga cctattgccc ccttcaggcg cgttaacttc ccattgcgtt gcagcaagac 11760
231 ttcgatttga ttgtcaagaa atcaaccaaa tacaactcac catatccgca qtttcggggc 11820
232 gaccttgagg caacgatgat gcgaataagg aactcaactg ggcacaaaat cgcgcaaaag 11880
233 attttcaata tttactcgct gctacaagac attcgaggtt ttcctctgct tgaatacagg 11940
234 gcaatagatg agtaagcatt taacaatttg ttaaaaccgt tgccttcgct cactgggacc 12000
235 ggcataagcc ggcctcttaa ccaaacgtta ggttaaccaag ggaattcac ttgagttgtt 12060
236 atgtattggg cacaacaacac cgcatttaaa ggacggtttt atagtaaat tcatcgagct 12120
237 gttgaactaa aatgcttata cgttttgctc tactacttgc tgttatgctc ctgcctgcat 12180
238 gctcgtcaaa gcaaaatcca acgcccgaat gtactgccag cgtccccccg cctctcttac 12240
239 ccgaaacatc cacagtatgc ctaggggaaa gatgtaattg ggaggtgcta tttccgtcag 12300
240 gaaaatacce tgcatccaca gaagcgtgca ggcgcctgtt ggtgcagaac cagcctctct 12360
241 cctaccgcgc agaagcactt gatcagtgta ttgaggggta cgttgggta gcggtttttc 12420
242 tgaatgccga cggggtccaa acatcagcaa aggtacttca atcatcgaat aaaattttcg 12480
243 acagaaatgc cttgctacag gccagtaata tattttttga gctatgaaa tgtcaatccg 12540
244 agcgttatga ttcggttggt ctgatgccat taaactaccg cactactccc tagtagcggg 12600
245 attgatccct acaaaattca ctacttacgt ccaagttgaa gtaggcagtt taacaactgg 12660
246 ttoagcgcgc tgccttcgct cactcgggac cgcctaaatt cggccctta ggcaaacgtt 12720
247 aactatcaga agggcggttg atgtcaagat ttgcgctcgc gttgattcac ggagtaccaa 12780
248 cgggttttct tgcatttgtt actttgtttg tctgtttcat ctacctaac cgattcgaga 12840
249 aagttggaag atactcaagc ggggtggggt ttgttggaag agttgtctgc gcattctatg 12900
250 ctatgggttt cgtgtccgca gttggccatc ttcttattga agcggcagtc aactgggggc 12960
251 tgcagcagct tggttatgag ctgcaaaact atgaaaaaag aaggacttgt agtagctgca 13020
252 agccgagcac tccagtgac tacatgttcg gcttgcctct cgggggtgtg cttggcgccg 13080
253 cctcggcaat ttggtctctg acgcgcctgg cgtccgata tgcgctgttt cgcggcgaaa 13140
254 actgatagct gaaccttcca tcgaggagat gcaaaagcgc tgcgtgcgcg catctacaaa 13200
255 gaccggaagc acctaatcca ggcgtctcca gcccgagcct gactggctgt ggctatcaac 13260

```

<210> 4  
<211> 42143  
<212> DNA  
<213> Unknown

<220>

<223> Description of Unknown Organism: Genomic DNA of  
Pseudomonas Alcaligenes NEB #585 (ATCC 55044)

<400> 4

```
atcgatcagc cagacttttc gcacacgggc ggaccttggg cgagtcagcg ctatgggttg 60
ccgctgtggg ttgtcagtgc ccgtacgcgc aatctgtttc tttcgcaggg catgtccggc 120
tgggcgttcc ggcccgttct ggtcaccgac tcggctctct atgagcgcta tctcgctcta 180
agtcaggaac tttgcgcact gcttcgtgat gcaccgcaga gcaagctcga agaccgtgat 240
tggtaaagcgg gggctattcg atcagtctcg gagcgaccaa actccagaaa cgacaaggcc 300
ctgaaaaaaa agcaggggctt cgtctttgcy ggccaatgga atcggacctc tttccgcctc 360
tgcatgtaac tggcttttgt ttgccaatc tgccatctc atgccggcca tgttggccag 420
tgccctgcac atttggcctt tggtttcgac actttttcga cagccctgct agacatccct 480
ccctctgccc tcgtaacttc tgttccgatg gtgtcgcttg gcactatggc cttgtcgagt 540
gtcgcttttc atccagccta atgccgcgat tgccctcgct agctgtagct gaatcaagga 600
cttagcggac gacaaggaat gttatgcgaa acatgtggcg gaataaatta cgccgcagt 660
ttcgtctact tatagttagg ctacatatga gaatcagcgc agaccagctt gctcaagaat 720
cactgactga gttcggcgct ctggcggtta agcttctggc aacgcgagag cttagccagt 780
tgtccgagaa gtttgggtat gcactggcct tcggaaggga accggcggtt gccatagctg 840
aggaccttgc taggtgcttg tgcggacaaa atgcttcgcc ggcatctgaa taccctaaaa 900
tcaccgttaa gtatttcaag gaaaacgaaa gtagtctgtt ggcaactcga gagtgttatg 960
tacaaatgac cgcaagcgca aacattcttt tagagctggt tgccgcacga aatggagagg 1020
caataaatct gtatctagaa ggcttgagtg ttgtagccta acaatgcgct caaagcgctc 1080
acttcgttcg ctgggaccgg cgaagccggc cccttagctt aatcgttaga aaccatcatg 1140
gataactggg acaacaccat cgaatacaca acccatgtag ccgaaaaact agaggcactt 1200
ggagaaacaa agtacgaccg cgaggcttat gaattcgcgc tagaggcata ccagtatgcy 1260
cctgaatatc atgaaaatat tcccacgcgc cctctcaatc ttgggctcgc gtaccatgta 1320
agcgccttca actttgcaca ctgctatgta cttcacgcta aagaagtgtt tgaagctcca 1380
aaagacacac tgagctcctg gggcgatttt tccctcaacg acatttgtga aattgtttat 1440
gggttagtcc gtattggctt gctggaccaa ggccccgaag acaaaaaaga gcagtttgaa 1500
gggttagttt taatcaccga cgtgctgtga tgtcttctaa ctactggttc aagtcgttcg 1560
cttcgctcac tcgggaccgg cttaaagcgg ccccttaacc aaacgttagc cacctcacga 1620
agatttgagg cccgcgtgaa caaagtcgat acaacaaaaa ttaaacgga ttttctggca 1680
cgaattgatg aaaaaagagc gtggtttgat cgtatggcta cgcttataag cgggacaaac 1740
accgagttaa ccgaccttaa ttttctttgc gagaactata taacatcaat atacgtagag 1800
ctcgaatgct taatatcaga tttatttcat ggctacataa ataacaaca caagacctac 1860
atggcgcaca ttcaatcaaa aatcaagaac tccataactg acaagtactc tgcatggcac 1920
gccacccata caacattcgc aggtccagag catattaatt cagcacagct cagcacgctc 1980
cttgatccaa caagctggaa catcacattt aaagacgttt ccgcaatgaa agtacgagca 2040
aaggaatacc tttcctcagt acacgaaaaa agattttcag gtatatctgc atccgatgga 2100
gctcttattg atgccgcaca tgcaatcaga aattgcattg cacacaacag cgaaagctcc 2160
agaaagggtt tgaacaccaa aattaaaagc ttaattacag gccagcttg ctcaaagtgc 2220
ggccttgaac tcaccacaaa tagtgtgacc aaaataggaa agtatctccg tgcaaagtgc 2280
cagcaaagca tgcgagtgtt gatttactca gatcgaataa aatctatcgg cctaagctta 2340
taagtgtggg ctaacaatgc gctcaactgt cgctcacttc gttcgttgga cagccaaaag 2400
ctacgctttt gtctgcccgt tagcttaatc gttaggaggc tctgcatgac tcgtgcaaca 2460
gacaggttcg aagagcttct gcaatcacat gagtctctag ggcatattat tcgttgggtt 2520
gcgatattcg aaggccgtct tgacggtgtg ttatcagttc attttctg acttgaaagc 2580
acctatgaat tctacgaact catactttcc aggttgtctt tctacgaaaa aattgaaatc 2640
ctgagaaaaa ttgatttttg taacagtctc aaatcccaag aaaatacagc gctgcacctc 2700
gacaaactga ggcgattgcy taacgcattg gcgcatgcag cacacatgcc acctgatgaa 2760
atcatgaagt tgtgctctga taagtggata gagtcccttg tgctcggaata tccaaagtcc 2820
```

09/70/626

6

see p. 7

09/20/2026 2

gctcggcaat	ttggetctgg	acgcgcctgg	cgctcgcgata	tgctcgtgttt	cgcgggcgaaa	13140
actgatagct	gaacctttcca	tcgaggagat	gcaaaaagcgc	tgctcgcgcgc	catctacaaa	13200
gaccggaagc	acctcatcca	ggcgcctctca	gcccagagcct	gactggctgt	ggctatcaac	13260
acctcttcga	taccactacc	cgccagaaac	gacaaaagccc	tgcaaaaagc	agggctttgt	13320
ctttggggat	ctggagcggg	cgaagggaat	cgaaccctcg	tcatgagctt	gggaagctca	13380
ggtaatgcca	ttatacgacg	cccgcctcggg	cggtcgactt	tttaccagaa	tcgcccggga	13440
aggtgaagcc	gggcgcgcgt	cttgcgccccg	ttttattgcc	gggcgcctca	tagcgccacg	13500
gcccgtggct	ctcgttccac	gctgcgtgcg	tggccctgcg	tgggtgccag	caggaaggcc	13560
agcagggcac	cgcgggctctg	catccaggcg	gccttgtgtt	ccatgtcgag	gaagtggccg	13620
gcctgggcga	tgttgcgga	ctcgcagtg	cgcacgtact	gggtgaacag	gcgcgcgtca	13680
gcccgggtgg	tgtactcgtc	ccactcgccg	ttgacgaaca	gcagcgggat	ctcgatctgc	13740
ccggcgaaac	tgacgcagga	gcgcccgcgc	ttgttcagca	cggtttccac	gtggtgactc	13800
atttgctcat	attcatagcg	ctccaggccg	gtgacgtgtc	gatggttgta	gcgcttgaac	13860
agcaggggca	ggtgcttgcc	gatggtgccc	ttgagcacca	tgccgatgct	ctcgcggtcg	13920
cactcgcgca	tcaccaccag	gcccggcgcgc	aggtagccga	gcatggcgct	gttgacgac	13980
ggcgagaagg	agttgatcac	cgcacgctcg	atccgcgatg	gacgcggggc	cagcgcctgg	14040
aggggtggcga	tgccgccccca	ggagaaagga	cagcacgctg	ttcgaggggc	aaaatgttcg	14100
accagctcca	ggaagatgtc	ggcttctttc	ctcgcsagctg	aaggrbscna	agcttctggt	14160
acgaacctgg	gggcgcctccg	gcacgcacaa	gggcatcgac	atcttcgccc	gccagggcac	14220
cccgggtgctc	gccccagct	acggcatcgt	ggtgtttcgc	gacgagctcg	acatgggcgg	14280
caagggtactg	ctgatgctcg	gccccaaatg	gcgcctgcac	tacttcgccc	acctcgacag	14340
ctacagcgcc	ctgcccggcc	aaccggtact	tcccggcgcc	ccactcggca	cggtaggcag	14400
caccggcaac	gcccagggca	agccgccccca	tctgcactac	tgcagctgca	ccctgttgcc	14460
ctatccctgg	cgctgggaca	acagcactca	gggctggaag	aaaatgttct	acctcgaccc	14520
cacgccaatg	ctgaacgaag	cggcagtaga	cagccgaaaa	accagccagt	agcgtcgcag	14580
gggaatgcac	caccggtctt	gcccgatccg	cctgtccttt	taccaatcgc	agaagagtcg	14640
cttttgtcga	atcgccctgtg	aggaaaaaca	aggacttgct	ggacgacaag	gaacgttatg	14700
cgacacaagt	ggcggaataa	attacgccat	ttgtgtcgtc	tacttatagt	tatatgctga	14760
tctagatatg	aagtacaaaa	acataaaatc	agcaatccac	aatttcgggc	acagctttgt	14820
aagctcagtg	aactatgttg	accatgattt	cgttgccgac	gaaattggga	agattcacia	14880
gaaaggctat	gatattgaaa	taaactggct	tacaaggag	ttcaagccc	ctcagcttga	14940
gtcagagaga	ataaaaaaat	caattggtta	ttgggggtgac	aacctaaaga	aacattgtgc	15000

7 see  
item 10  
on Enn  
summary  
sheet

due to size of sequence 4, only the first  
error in that sequence shown. Identical errors  
appear elsewhere in the sequence.

FYI:

**Please Note:**

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/701,626

DATE: 01/30/2001

TIME: 11:15:34

Input Set : A:\Neb-165a.app

Output Set: N:\CRF3\01302001\I701626.raw

L:12 M:270 C: Current Application Number differs, Replaced Current Application Number  
L:13 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:897 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:4  
L:897 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:4  
L:897 M:340 W: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:4  
L:1019 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:4  
L:1019 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:4  
M:340 Repeated in SeqNo=4  
L:1141 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:4  
L:1141 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:4